

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANTS: Seed, Brian et al.
- (ii) TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected Cells by  
Chimeric CD4 Receptor-Bearing Cells
- (iii) NUMBER OF SEQUENCES: 27
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Fish & Richardson
  - (B) STREET: 225 Franklin Street
  - (C) CITY: Boston
  - (D) STATE: MA
  - (E) COUNTRY: USA
  - (F) ZIP: 02110-2804
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
  - (B) COMPUTER: IBM PS/2 Model 50Z or 55SX
  - (C) OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
  - (D) SOFTWARE: Wordperfect (Version 5.0)
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 07/847,566
  - (B) FILING DATE: March 6, 1992
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 07/665,961
  - (B) FILING DATE: March 7, 1991
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Clark, Paul T.
  - (B) REGISTRATION NUMBER: 30,162
  - (C) REFERENCE/DOCKET NUMBER: 00786/212001
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (617) 542-5070
  - (B) TELEFAX: (617) 542-8906
  - (C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1728 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

009957 0844

ATGAACCGGG GAGTCCCTTT TAGGCACTTG CTTCTGGTGC TGCAACTGGC	50
GCTCCTCCCA GCAGCCACTC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG	100
GGGATACAGT GGAAGTGACC TGTACAGCTT CCCAGAAGAA GAGCATACAA	150
TTCCACTGGA AAAACTCCAA CCAGATAAAG ATTCTGGGAA ATCAGGGCTC	200
CTTCTTAACT AAAGGTCCAT CCAAGCTGAA TGATCGCGCT GACTCAAGAA	250
GAAGCCTTTG GGACCAAGGA AACTTCCCCC TGATCATCAA GAATCTTAAG	300
ATAGAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA	350
GGTGCAATTG CTAGTGTTTG GATTGACTGC CAACTCTGAC ACCCACCTGC	400
TTCAGGGGCA GAGCCTGACC CTGACCTTGG AGAGCCCCCC TGGTAGTAGC	450
CCCTCAGTGC AATGTAGGAG TCCAAGGGGT AAAAACATAC AGGGGGGGAA	500
GACCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC ACCTGGACAT	550
GCACTGTCTT GCAGAACCAG AAGAAGGTGG AGTTCAAAAT AGACATCGTG	600
GTGCTAGCTT TCCAGAAGGC CTCCAGCATA GTCTATAAGA AAGAGGGGGA	650
ACAGGTGGAG TTCTCCTTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG	700
GCAGTGGCGA GCTGTGGTGG CAGGCGGAGA GGGCTTCCTC CTCCAAGTCT	750
TGGATCACCT TTGACCTGAA GAACAAGGAA GTGTCTGTAA AACGGGTTAC	800
CCAGGACCCT AAGCTCCAGA TGGGCAAGAA GCTCCCGCTC CACCTCACCC	850
TGCCCCAGGC CTTGCCCTCAG TATGCTGGCT CTGGAAACCT CACCCTGGCC	900
CTTGAAGCGA AAACAGGAAA GTTGCAATCAG GAAGTGAACC TGGTGGTGAT	950
GAGAGCCACT CAGCTCCAGA AAAATTGAC CTGTGAGGTG TGGGGACCCA	1000
CCTCCCCTAA GCTGATGCTG AGCTTGAAAC TGGAGAACAA GGAGGCAAAG	1050
GTCTCGAAGC GGGAGAAGCC GGTGTGGGTG CTGAACCCTG AGGCGGGGAT	1100
GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCTGCTG GAATCCAACA	1150
TCAAGGTTCT GCCCACATGG TCCACCCCGG TGCACGCGGA TCCCAAATC	1200
TGCTACTTGC TAGATGGAAT CCTCTTCATC TACGGAGTCA TCATCACAGC	1250
CCTGTACCTG AGAGCAAAAT TCAGCAGGAG TGCAGAGACT GCTGCCAACC	1300
TGCAGGACCC CAACCAGCTC TACAATGAGC TCAATCTAGG GCGAAGAGAG	1350
GAATATGACG TCTTGAGAA GAAGCGGGCT CGGGATCCAG AGATGGGAGG	1400
CAACACAGCAG AGGAGGAGGA ACCCCCAGGA AGGCGTATAC AATGCACTGC	1450
AGAAAGACAA GATGCCAGAA GCCTACAGTG AGATCGGCAC AAAAGGCGAG	1500

AGGCCGAGAG GCAAGGGGCA CGATGGCCTT TACCAGGACA GCCACTTCCA	1550
AGCAGTGCAG TTCGGGAACA GAAGAGAGAG AGAAGGTTCA GAACTCACAA	1600
GGACCCCTTG GTTAAGAGCC CGCCCCAAG GTGAAAGCAC CCAGCAGAGT	1650
AGCCAATCCT GTGCCAGCGT CTTCAGCATC CCCACTCTGT GGAGTCCATG	1700
GCCACCCAGT AGCAGCTCCC AGCTCTAA	1728

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1389 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGAACCGGG GAGTCCCTTT TAGGCACTTG CTTCTGGTGC TGCAACTGGC	50
GCTCCTCCCA GCAGCCACTC AGGGAAACAA AGTGGTGCTG GGCAAAAAG	100
GGGATACAGT GGAAGTGACC TGTACAGCTT CCCAGAAGAA GAGCATACAA	150
TTCCACTGGA AAAACTCCAA CCAGATAAAG ATTCTGGGAA ATCAGGGCTC	200
CTTCTTAACT AAAGGTCCAT CCAAGCTGAA TGATCGCGCT GACTCAAGAA	250
GAAGCCTTTG GGACCAAGGA AACTTCCCCC TGATCATCAA GAATCTTAAG	300
ATAGAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA	350
GGTGCAATTG CTAGTGTTCG GATTGACTGC CAACTCTGAC ACCCACCTGC	400
TTCAGGGGCA GAGCCTGACC CTGACCTTGG AGAGCCCCCC TGGTAGTAGC	450
CCCTCAGTGC AATGTAGGAG TCCAAGGGGT AAAACATAC AGGGGGGGAA	500
GACCCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC ACCTGGACAT	550
GCACTGTCTT GCAGAACCAG AAGAAGGTGG AGTTCAAAAT AGACATCGTG	600
GTGCTAGCTT TCCAGAAGGC CTCCAGCATA GTCTATAAGA AAGAGGGGGA	650
ACAGGTGGAG TTCTCCTTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG	700
GCAGTGGCGA GCTGTGGTGG CAGGCGGAGA GGGCTTCCTC CTCCAAGTCT	750
TGGATCACCT TTGACCTGAA GAACAAGGAA GTGTCTGTAA AACGGGTTAC	800
CCAGGACCCT AAGCTCCAGA TGGGCAAGAA GCTCCCGCTC CACCTCACCC	850
TGCCCCAGGC CTTGCCTCAG TATGCTGGCT CTGGAAACCT CACCCTGGCC	900
CTTGAAGCGA AAACAGGAAA GTTGCATCAG GAAGTGAACC TGGTGGTGAT	950

F04280-4656560

GAGAGCCACT CAGCTCCAGA AAAATTGAC CTGTGAGGTG TGGGGACCCA	1000
CCTCCCCTAA GCTGATGCTG AGCTTGAAC TGGAGAACAA GGAGGCAAAG	1050
GTCTCGAAGC GGGAGAAGCC GGTGTGGGTG CTGAACCCTG AGGCGGGGAT	1100
GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCTGTCTG GAATCCAACA	1150
TCAAGGTTCT GCCCACATGG TCCACCCCGG TGCACGCGGA TCCGCAGCTC	1200
TGCTATATCC TGGATGCCAT CCTGTTTTTG TATGGTATTG TCCTTACCCT	1250
GCTCTACTGT CGACTCAAGA TCCAGGTCCG AAAGGCAGAC ATAGCCAGCC	1300
GTGAGAAATC AGATGCTGTC TACACGGGCC TGAACACCCG GAACCAGGAG	1350
ACATATGAGA CTCTGAAACA TGAGAAACCA CCCCAATAG	1389

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1599 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAACCGGG GAGTCCCTTT TAGGCACTTG CTTCTGGTGC TGCAACTGGC	50
GCTCCTCCCA GCAGCCACTC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG	100
GGGATACAGT GGAAC TGACCTG TGTACAGCTT CCCAGAAGAA GAGCATAACA	150
TTCCACTGGA AAAACTCCAA CCAGATAAAG ATTCTGGGAA ATCAGGGCTC	200
CTTCTTAACT AAAGGTCCAT CCAAGCTGAA TGATCGCGCT GACTCAAGAA	250
GAAGCCTTTG GGACCAAGGA AACTTCCCCC TGATCATCAA GAATCTTAAG	300
ATAGAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA	350
GGTGCAATTG CTAGTGTTTG GATTGACTGC CAACTCTGAC ACCCACCTGC	400
TTCAGGGGCA GAGCCTGACC CTGACCTTGG AGAGCCCCCC TGGTAGTAGC	450
CCCTCAGTGC AATGTAGGAG TCCAAGGGGT AAAACATAC AGGGGGGGAA	500
GACCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC ACCTGGACAT	550
GCACTGTCTT GCAGAACCAG AAGAAGGTGG AGTTCAAAAT AGACATCGTG	600
GTGCTAGCTT TCCAGAAGGC CTCCAGCATA GTCTATAAGA AAGAGGGGGA	650
ACAGGTGGAG TTCTCCTTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG	700
GCAGTGGCGA GCTGTGGTGG CAGGCGGAGA GGGCTTCCTC CTCCAAGTCT	750

F01230-255660

TGGATCACCT TTGACCTGAA GAACAAGGAA GTGTCTGTAA AACGGGTTAC	800
CCAGGACCCT AAGCTCCAGA TGGGCAAGAA GCTCCCGCTC CACCTCACCC	850
TGCCCCAGGC CTTCCTCAG TATGCTGGCT CTGGAAACCT CACCCTGGCC	900
CTTGAAGCGA AAACAGGAAA GTTGCAACAG GAAGTGAACC TGGTGGTGAT	950
GAGAGCCACT CAGCTCCAGA AAAATTTGAC CTGTGAGGTG TGGGGACCCA	1000
CCTCCCCCTAA GCTGATGCTG AGCTTGAAAC TGGAGAACAA GGAGGCAAAG	1050
GTCTCGAAGC GGGAGAAGCC GGTGTGGGTG CTGAACCCTG AGGCGGGGAT	1100
GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCTGCTG GAATCCAACA	1150
TCAAGGTTCT GCCCACATGG TCCACCCCGG TGCACGCGGA TCCCAAACCTC	1200
TGCTACCTGC TGGATGGAAT CCTCTTCATC TATGGTGTC TTTCTACTGC	1250
CTTGTTCTCTG AGAGTGAAGT TCAGCAGGAG CGCAGAGCCC CCCGCGTACC	1300
AGCAGGGCCA GAACCAGCTC TATAACGAGC TCAATCTAGG ACGAAGAGAG	1350
GAGTACGATG TTTTGGACAA GAGACGTGGC CGGGACCCTG AGATGGGGGG	1400
AAAGCCGAGA AGGAAGAACC CTCAGGAAGG CCTGTACAAT GAACTGCAGA	1450
AAGATAAGAT GCGGGAGGCC TACAGTGAGA TTGGGATGAA AGGCGAGCGC	1500
CGGAGGGGCA AGGGGCACGA TGGCCTTTAC CAGGGTCTCA GTACAGCCAC	1550
CAAGGACACC TACGACGCCC TTCACATGCA GGCCCTGCCC CCTCGCTAA	1599

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 575 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Asn	Arg	Gly	Val	Pro	Phe	Arg	His	Leu	Leu	Leu	Val	Leu	Gln	Leu
1				5					10					15	
Ala	Leu	Leu	Pro	Ala	Ala	Thr	Gln	Gly	Asn	Lys	Val	Val	Leu	Gly	Lys
			20					25					30		
Lys	Gly	Asp	Thr	Val	Glu	Leu	Thr	Cys	Thr	Ala	Ser	Gln	Lys	Lys	Ser
		35					40					45			
Ile	Gln	Phe	His	Trp	Lys	Asn	Ser	Asn	Gln	Ile	Lys	Ile	Leu	Gly	Asn
	50				55				60						
Gln	Gly	Ser	Phe	Leu	Thr	Lys	Gly	Pro	Ser	Lys	Leu	Asn	Asp	Arg	Ala
	65			70				75						80	
Asp	Ser	Arg	Arg	Ser	Leu	Trp	Asp	Gln	Gly	Asn	Phe	Pro	Leu	Ile	Ile
			85				90						95		
Lys	Asn	Leu	Lys	Ile	Glu	Asp	Ser	Asp	Thr	Tyr	Ile	Cys	Glu	Val	Gln
		100				105						110			
Asp	Gln	Lys	Glu	Glu	Val	Gln	Leu	Leu	Val	Phe	Gly	Leu	Thr	Ala	Asn
	115					120						125			

Ser	Asp	Thr	His	Leu	Leu	Gln	Gly	Gln	Ser	Leu	Thr	Leu	Thr	Leu	Glu
130						135						140			
Ser	Pro	Pro	Gly	Ser	Ser	Pro	Ser	Val	Gln	Cys	Arg	Ser	Pro	Arg	Gly
145					150					155					160
Lys	Asn	Ile	Gln	Gly	Gly	Lys	Thr	Leu	Ser	Val	Ser	Gln	Leu	Glu	Leu
				165					170						175
Gln	Asp	Ser	Gly	Thr	Trp	Thr	Cys	Thr	Val	Leu	Gln	Asn	Gln	Lys	Lys
			180					185					190		
Val	Glu	Phe	Lys	Ile	Asp	Ile	Val	Val	Leu	Ala	Phe	Gln	Lys	Ala	Ser
		195					200					205			
Ser	Ile	Val	Tyr	Lys	Lys	Glu	Gly	Glu	Gln	Val	Glu	Phe	Ser	Phe	Pro
	210					215					220				
Leu	Ala	Phe	Thr	Val	Glu	Lys	Leu	Thr	Gly	Ser	Gly	Glu	Leu	Trp	Trp
225					230					235					240
Gln	Ala	Glu	Arg	Ala	Ser	Ser	Ser	Lys	Ser	Trp	Ile	Thr	Phe	Asp	Leu
				245					250					255	
Lys	Asn	Lys	Glu	Val	Ser	Val	Lys	Arg	Val	Thr	Gln	Asp	Pro	Lys	Leu
			260					265					270		
Gln	Met	Gly	Lys	Lys	Leu	Pro	Leu	His	Leu	Thr	Leu	Pro	Gln	Ala	Leu
		275					280					285			
Pro	Gln	Tyr	Ala	Gly	Ser	Gly	Asn	Leu	Thr	Leu	Ala	Leu	Glu	Ala	Lys
	290					295					300				
Thr	Gly	Lys	Leu	His	Gln	Glu	Val	Asn	Leu	Val	Val	Met	Arg	Ala	Thr
305					310					315					320
Gln	Leu	Gln	Lys	Asn	Leu	Thr	Cys	Glu	Val	Trp	Gly	Pro	Thr	Ser	Pro
				325					330					335	
Lys	Leu	Met	Leu	Ser	Leu	Lys	Leu	Glu	Asn	Lys	Glu	Ala	Lys	Val	Ser
			340					345					350		
Lys	Arg	Glu	Lys	Pro	Val	Trp	Val	Leu	Asn	Pro	Glu	Ala	Gly	Met	Trp
		355					360					365			
Gln	Cys	Leu	Leu	Ser	Asp	Ser	Gly	Gln	Val	Leu	Leu	Glu	Ser	Asn	Ile
	370					375						380			
Lys	Val	Leu	Pro	Thr	Trp	Ser	Thr	Pro	Val	His	Ala	Asp	Pro	Lys	Leu
385					390					395					400
Cys	Tyr	Leu	Leu	Asp	Gly	Ile	Leu	Phe	Ile	Tyr	Gly	Val	Ile	Ile	Thr
				405					410					415	
Ala	Leu	Tyr	Leu	Arg	Ala	Lys	Phe	Ser	Arg	Ser	Ala	Glu	Thr	Ala	Ala
			420					425					430		
Asn	Leu	Gln	Asp	Pro	Asn	Gln	Leu	Tyr	Asn	Glu	Leu	Asn	Leu	Gly	Arg
		435					440					445			
Arg	Glu	Glu	Tyr	Asp	Val	Leu	Glu	Lys	Lys	Arg	Ala	Arg	Asp	Pro	Glu
	450					455					460				
Met	Gly	Gly	Lys	Gln	Gln	Arg	Arg	Arg	Asn	Pro	Gln	Glu	Gly	Val	Tyr
465					470					475					480
Asn	Ala	Leu	Gln	Lys	Asp	Lys	Met	Pro	Glu	Ala	Tyr	Ser	Glu	Ile	Gly
			485						490					495	
Thr	Lys	Gly	Glu	Arg	Arg	Arg	Gly	Lys	Gly	His	Asp	Gly	Leu	Tyr	Gln
			500					505					510		
Asp	Ser	His	Phe	Gln	Ala	Val	Gln	Phe	Gly	Asn	Arg	Arg	Glu	Arg	Glu
		515					520						525		
Gly	Ser	Glu	Leu	Thr	Arg	Thr	Leu	Gly	Leu	Arg	Ala	Arg	Pro	Lys	Gly
	530					535					540				
Glu	Ser	Thr	Gln	Gln	Ser	Ser	Gln	Ser	Cys	Ala	Ser	Val	Phe	Ser	Ile
555					550					565					560
Pro	Thr	Leu	Trp	Ser	Pro	Trp	Pro	Pro	Ser	Ser	Ser	Ser	Gln	Leu	
				565					570					575	

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 462 amino acids

(ii) MOLECULE TYPE: protein

Met	Asn	Arg	Gly	Val	Pro	Phe	Arg	His	Leu	Leu	Val	Leu	Gln	Leu
1				5					10				15	
Ala	Leu	Leu	Pro	Ala	Ala	Thr	Gln	Gly	Asn	Lys	Val	Val	Leu	Gly
			20					25					30	Lys
Lys	Gly	Asp	Thr	Val	Glu	Leu	Thr	Cys	Thr	Ala	Ser	Gln	Lys	Lys
		35					40					45		Ser
Ile	Gln	Phe	His	Trp	Lys	Asn	Ser	Asn	Gln	Ile	Lys	Ile	Leu	Gly
	50				55						60			Asn
Gln	Gly	Ser	Phe	Leu	Thr	Lys	Gly	Pro	Ser	Lys	Leu	Asn	Asp	Arg
	65				70					75				Ala
Asp	Ser	Arg	Arg	Ser	Leu	Trp	Asp	Gln	Gly	Asn	Phe	Pro	Leu	Ile
			85						90					95
Lys	Asn	Leu	Lys	Ile	Glu	Asp	Ser	Asp	Thr	Tyr	Ile	Cys	Glu	Val
			100					105					110	Glu
Asp	Gln	Lys	Glu	Glu	Val	Gln	Leu	Leu	Val	Phe	Gly	Leu	Thr	Ala
			115				120					125		Asn
Ser	Asp	Thr	His	Leu	Leu	Gln	Gly	Gln	Ser	Leu	Thr	Leu	Thr	Leu
	130					135					140			Glu
Ser	Pro	Pro	Gly	Ser	Ser	Pro	Ser	Val	Gln	Cys	Arg	Ser	Pro	Arg
	145				150					155				Gly
Lys	Asn	Ile	Gln	Gly	Gly	Lys	Thr	Leu	Ser	Val	Ser	Gln	Leu	Glu
			165						170					Leu
Gln	Asp	Ser	Gly	Thr	Trp	Thr	Cys	Thr	Val	Leu	Gln	Asn	Gln	Lys
			180					185					190	Lys
Val	Glu	Phe	Lys	Ile	Asp	Ile	Val	Val	Leu	Ala	Phe	Gln	Lys	Ala
		195					200					205		Ser
Ser	Ile	Val	Tyr	Lys	Lys	Glu	Gly	Glu	Gln	Val	Glu	Phe	Ser	Phe
	210					215					220			Pro
Leu	Ala	Phe	Thr	Val	Glu	Lys	Leu	Thr	Gly	Ser	Gly	Glu	Leu	Trp
	225				230					235				Trp
Gln	Ala	Glu	Arg	Ala	Ser	Ser	Ser	Lys	Ser	Trp	Ile	Thr	Phe	Asp
			245						250					Leu
Lys	Asn	Lys	Glu	Val	Ser	Val	Lys	Arg	Val	Thr	Gln	Asp	Pro	Lys
			260					265					270	Leu
Gln	Met	Gly	Lys	Lys	Leu	Pro	Leu	His	Leu	Thr	Leu	Pro	Gln	Ala
		275					280					285		Leu
Pro	Gln	Tyr	Ala	Gly	Ser	Gly	Asn	Leu	Thr	Leu	Ala	Leu	Glu	Ala
		290				295					300			Lys
Thr	Gly	Lys	Leu	His	Gln	Glu	Val	Asn	Leu	Val	Val	Met	Arg	Ala
				310						315				Thr
Gln	Leu	Gln	Lys	Asn	Leu	Thr	Cys	Glu	Val	Trp	Gly	Pro	Thr	Ser
			325						330					Pro
Lys	Leu	Met	Leu	Ser	Leu	Lys	Leu	Glu	Asn	Lys	Glu	Ala	Lys	Val
			340					345					350	Ser
Lys	Arg	Glu	Lys	Pro	Val	Trp	Val	Leu	Asn					

(2) INFORMATION FOR SEQ ID NO:6:

(A) LENGTH: 532 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Asn	Arg	Gly	Val	Pro	Phe	Arg	His	Leu	Leu	Leu	Val	Leu	Gln	Leu
1				5					10					15	
Ala	Leu	Leu	Pro	Ala	Ala	Thr	Gln	Gly	Asn	Lys	Val	Val	Leu	Gly	Lys
			20					25					30		
Lys	Gly	Asp	Thr	Val	Glu	Leu	Thr	Cys	Thr	Ala	Ser	Gln	Lys	Lys	Ser
		35					40					45			
Ile	Gln	Phe	His	Trp	Lys	Asn	Ser	Asn	Gln	Ile	Lys	Ile	Leu	Gly	Asn
	50				55						60				
Gln	Gly	Ser	Phe	Leu	Thr	Lys	Gly	Pro	Ser	Lys	Leu	Asn	Asp	Arg	Ala
	65			70						75				80	
Asp	Ser	Arg	Arg	Ser	Leu	Trp	Asp	Gln	Gly	Asn	Phe	Pro	Leu	Ile	Ile
			85					90						95	
Lys	Asn	Leu	Lys	Ile	Glu	Asp	Ser	Asp	Thr	Tyr	Ile	Cys	Glu	Val	Glu
			100					105					110		
Asp	Gln	Lys	Glu	Glu	Val	Gln	Leu	Leu	Val	Phe	Gly	Leu	Thr	Ala	Asn
		115					120					125			
Ser	Asp	Thr	His	Leu	Leu	Gln	Gly	Gln	Ser	Leu	Thr	Leu	Thr	Leu	Glu
	130					135					140				
Ser	Pro	Pro	Gly	Ser	Ser	Pro	Ser	Val	Gln	Cys	Arg	Ser	Pro	Arg	Gly
	145			150						155				160	
Lys	Asn	Ile	Gln	Gly	Gly	Lys	Thr	Leu	Ser	Val	Ser	Gln	Leu	Glu	Leu
			165					170						175	
Gln	Asp	Ser	Gly	Thr	Trp	Thr	Cys	Thr	Val	Leu	Gln	Asn	Gln	Lys	Lys
		180						185					190		
Val	Glu	Phe	Lys	Ile	Asp	Ile	Val	Val	Leu	Ala	Phe	Gln	Lys	Ala	Ser
		195					200					205			
Ser	Ile	Val	Tyr	Lys	Lys	Glu	Gly	Gln	Gln	Val	Glu	Phe	Ser	Phe	Pro
	210					215					220				
Leu	Ala	Phe	Thr	Val	Glu	Lys	Leu	Thr	Gly	Ser	Gly	Glu	Leu	Trp	Trp
	225				230					235				240	
Gln	Ala	Glu	Arg	Ala	Ser	Ser	Ser	Lys	Ser	Trp	Ile	Thr	Phe	Asp	Leu
			245					250						255	
Lys	Asn	Lys	Glu	Val	Ser	Val	Lys	Arg	Val	Thr	Gln	Asp	Pro	Lys	Leu
			260					265					270		
Gln	Met	Gly	Lys	Lys	Leu	Pro	Leu	His	Leu	Thr	Leu	Pro	Gln	Ala	Leu
		275					280					285			
Pro	Gln	Tyr	Ala	Gly	Ser	Gly	Asn	Leu	Thr	Leu	Ala	Leu	Glu	Ala	Lys
	290					295					300				
Thr	Gly	Lys	Leu	His	Gln	Glu	Val	Asn	Leu	Val	Val	Met	Arg	Ala	Thr
				310						315					320
Gln	Leu	Gln	Lys	Asn	Leu	Thr	Cys	Glu	Val	Trp	Gly	Pro	Thr	Ser	Pro
				325					330					335	
Lys	Leu	Met	Leu	Ser	Leu	Lys									



Gln Cys Leu Leu Ser Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile  
 370 375 380  
 Lys Val Leu Pro Thr Trp Ser Thr Pro Val His Ala Asp Pro Lys Leu  
 385 390 395 400  
 Cys Tyr Leu Leu Asp Gly Ile Leu Phe Ile Tyr Gly Val Ile Leu Thr  
 405 410 415  
 Ala Leu Phe Leu Arg Val Lys Phe Ser Arg Ser Ala Glu Pro Pro Ala  
 420 425 430  
 Tyr Gln Gln Gly Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg  
 435 440 445  
 Arg Glu Glu Tyr Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu  
 450 455 460  
 Met Gly Gly Lys Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn  
 465 470 475 480  
 Glu Leu Gln Lys Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met  
 485 490 495  
 Lys Gly Glu Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly  
 500 505 510  
 Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala  
 515 520 525  
 Leu Pro Pro Arg  
 530

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCGGGGTGA CCGTCCCCTC CAGCAGCTTG GGC

33

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCGGGGATC CGTCGTCCAG AGCCCGTCCA GCTCCCCGTC CTGGGCCTCA

50

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGCGGGGAGC TCGTTATAGA GCTGGTTTGC CGCCGAATTC TTATCCCG

48

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 bases  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGCGGGGCGG CCACGCGTCC TCGCCAGCAC ACA

33

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 bases  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGCGGGACGC GTTTCAGCCG TCCTCGCCAG CACACA

36

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 bases  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGCGGGACGC GTGACCCTGA GATGGGGGGA AAG

33

(2) INFORMATION FOR SEQ ID NO:17:

(A) LENGTH: 33 bases  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

33

(A) LENGTH: 26 bases  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

26

(A) LENGTH: 42 bases  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

42

(A) LENGTH: 30 bases  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

30

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCGGGGGCGG CCGCCTAAAT ACGGTTCTGG TC

32

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TCAGAAAGAG ACAACCTGAA GAAACCAACA A

31

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

10-1-80 15:56:50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TTGTTGGTTT CTTCAGGTTG TGTCTTCTG A

31

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 171 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met	Glu	His	Ser	Thr	Phe	Leu	Ser	Gly	Leu	Val	Leu	Ala	Thr	Leu	Leu
				5					10					15	
Ser	Gln	Val	Ser	Pro	Phe	Lys	Ile	Pro	Ile	Glu	Glu	Leu	Glu	Asp	Arg
			20					25					30		
Val	Phe	Val	Asn	Cys	Asn	Thr	Ser	Ile	Thr	Trp	Val	Glu	Gly	Thr	Val
			35				40					45			
Gly	Thr	Leu	Leu	Ser	Asp	Ile	Thr	Arg	Leu	Asp	Leu	Gly	Lys	Arg	Ile
			50			55					60				
Leu	Asp	Pro	Arg	Gly	Ile	Tyr	Arg	Cys	Asn	Gly	Thr	Asp	Ile	Tyr	Lys
					70					75				80	
Asp	Lys	Glu	Ser	Thr	Val	Gln	Val	His	Tyr	Arg	Met	Cys	Gln	Ser	Cys
					85					90				95	
Val	Glu	Leu	Asp	Pro	Ala	Thr	Val	Ala	Gly	Ile	Ile	Val	Thr	Asp	Val
					100				105				110		
Ala	Ile	Thr	Leu	Leu	Leu	Ala	Leu	Gly	Val	Phe	Cys	Phe	Ala	Gly	His
							120					125			
Glu	Thr	Gly	Arg	Leu	Ser	Gly	Ala	Ala	Asp	Thr	Gln	Ala	Leu	Leu	Arg
							135					140			
Asn	Asp	Gln	Val	Tyr	Gln	Pro	Leu	Arg	Asp	Arg	Asp	Asp	Ala	Gln	Tyr
					150					155					160
Ser	His	Leu	Gly	Gly	Asn	Trp	Ala	Arg	Asn	Lys					
					165					170					

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 182 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Met	Glu	Gln	Gly	Lys	Gly	Leu	Ala	Val	Leu	Ile	Leu	Ala	Ile	Ile	Leu
				5					10					15	
Leu	Gln	Gly	Thr	Leu	Ala	Gln	Ser	Ile	Lys	Gly	Asn	His	Leu	Val	Lys
				20				25					30		
Val	Tyr	Asp	Tyr	Gln	Glu	Asp	Gly	Ser	Val	Leu	Leu	Thr	Cys	Asp	Ala
				35			40					45			
Glu	Ala	Lys	Asn	Ile	Thr	Trp	Phe	Lys	Asp	Gly	Lys	Met	Ile	Gly	Phe
				50			55				60				

(2) INFORMATION FOR SEQ ID NO:26:

(A) LENGTH: 220 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

(2) INFORMATION FOR SEQ ID NO:27:

## (ii) MOLECULE TYPE: amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

[illegible]